OPTIMAL NODE DENSITY FOR TWO-DIMENSIONAL SENSOR ARRAYS

Youngchul Sung[†], H. Vincent Poor and Heejung Yu

ABSTRACT

The problem of optimal node density for *ad hoc* sensor networks deployed for making inferences about two dimensional correlated random fields is considered. Using a symmetric first order conditional autoregressive Gauss-Markov random field model, large deviations results are used to characterize the asymptotic per-node information gained from the array. This result then allows an analysis of the node density that maximizes the information under an energy constraint, yielding insights into the trade-offs among the information, density and energy.

1. INTRODUCTION

We consider the design of wireless *ad hoc* sensor networks for making inferences about correlated random fields that can model various physical processes, such as temperature, humidity or the density of a certain gas, in a two-dimensional (2-D) space. In particular, we consider the optimal density problem for sensor networks deployed for statistical inference such as detection or reconstruction of the underlying field. From the information-theoretic perspective, statistical inference via sensor networks can be viewed as a problem of extracting information about an underlying physical process using networked sensor nodes that consume energy for both sensing and communication. Thus, the optimal density problem can be formulated as follows.

Problem 1 *Given a sensor network deployed on a fixed coverage area of size $2L \times 2L$ and with total available energy E, find the node density μ_n that maximizes the total information I_t obtainable from the network.*

To address this problem, we model the signal field as a 2-D Gauss-Markov random field (GMRF), and consider the Kullback-Leibler information (KLI) and mutual information (MI) [1] as ways of quantifying inferential performance. (The operational meaning of the KLI is given by its appearance as the error exponent of the miss probability of Neyman-Pearson detection of the signal field in sensor noise, whereas that of the MI is given by its role as a measure of uncertainty reduction.) Our approach to determine the total information obtainable from a sensor network is based on the large deviations principle (LDP). That is, for large networks, the total information is approximately given by the product of the number of sensors and the asymptotic per-node information, or the asymptotic information rate. (The units of these intensive quantities is thus nats/sample.) Although closed-form expressions for the

asymptotic per-node information are not available for general 2-D signals, for the conditional autoregression (CAR) model closed-form expressions for the asymptotic KLI and MI rates have been determined by the authors in [2]. Based on these expressions for asymptotic information rates and their properties, in the current paper we investigate the problem of optimal node density. It is seen that there exists a density maximizing the total information obtainable under an energy constraint. The optimal density is easily obtained numerically, and the behavior of the total information as a function of the density is explained.

1.1. Related Work

The issues of optimal sensor density and optimal sampling have been considered based on LDP in previous work (e.g., [3]). However, most work in this area is based on one-dimensional (1-D) signal or time series models that do not capture the two-dimensionality of actual spatial signals. In contrast, our work is based on the LDP results obtained in [2], where a closed-form expression for the asymptotic KLI rate is obtained in the spectral domain. For a 2-D setting, an error exponent was obtained for the detection of 2-D GMRFs in [4], where the sensors are located randomly and the Markov graph is based on the nearest neighbor dependency enabling a loop-free graph. In that work, however, measurement noise was not considered, unlike the present analysis.

2. SIGNAL MODEL AND BACKGROUND

In this section, we briefly introduce our previous work [2] relevant to the sensor density problem. To simplify the problem and gain insight into the 2-D case, we assume that sensors are located on a 2-D lattice $\mathcal{I}_n = [-n : 1 : n]^2$, as shown in Fig. 1, and thus form a 2-D array. We model the underlying physical process as a 2-D GMRF and assume that each sensor has Gaussian measurement noise. So, the observation Y_{ij} of Sensor ij on the 2-D lattice $\mathcal{I}_n$ is given by

$$Y_{ij} = X_{ij} + W_{ij}, \quad ij \in \mathcal{I}_n, \tag{1}$$

where $\{W_{ij}\}$ represents independent and identically distributed (i.i.d.) zero-mean Gaussian measurement noise with variance σ^2, and $\{X_{ij}\}$ is a GMRF on $\mathcal{I}_n$, independent of $\{W_{ij}\}$. Note that the observation samples form a 2-D hidden GMRF on $\mathcal{I}_n$. In the following, we summarize our relevant LDP results on GMRFs that will be useful in the sequel.
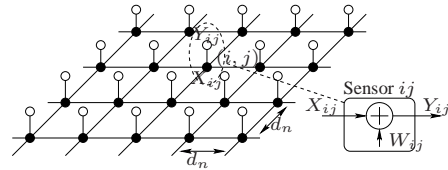

Fig. 1. 2-D sensor array on a lattice $\mathcal{I}_n$: Hidden Markov structure

Definition 1 (GMRF [5]) *A random vector* $\mathbf{X} = (X_1, X_2, \cdots, X_n) \in \mathbb{R}^n$ *is a Gauss-Markov random field with respect to (w.r.t.). a labelled graph* $\mathcal{G} = (\nu, \mathcal{E})$ *with mean vector* $\boldsymbol{\mu}$ *and precision matrix* $\mathbf{Q} > 0$, *if its probability density function is given by*

$$p(\mathbf{X}) = (2\pi)^{-n/2}|\mathbf{Q}|^{1/2} \exp\left(-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T \mathbf{Q}(\mathbf{X}-\boldsymbol{\mu})\right), \quad (2)$$

and $Q_{lm} \neq 0 \iff \{l, m\} \in \mathcal{E}$ *for all* $l \neq m$. *Here,* ν *is the set of all nodes* $\{1, 2, \cdots, n\}$ *and* $\mathcal{E}$ *is the set of edges connecting pairs of nodes, which represent the conditional dependence structure.*

Note that the 2-D indexing scheme ij in (1) can be properly converted to an 1-D scheme to apply Definition 1. From here on, we use the 2-D indexing scheme for convenience.

Definition 2 (The Conditional Autoregression (CAR)) *A GMRF* $\{X_{ij}\}$ *is said to be a conditional autoregression if it is specified using a set of full conditional normal distributions with means and precisions:*

$$\mathbb{E}\{X_{ij}|\mathbf{X}_{-ij}\} = -\frac{1}{\theta_{00}}\sum_{i'j' \in \mathcal{I}_\infty \neq 00} \theta_{i'j'} X_{i+i',j+j'}, \quad (3)$$

$$Prec\{X_{ij}|\mathbf{X}_{-ij}\} = \theta_{00} > 0, \quad (4)$$

where $\mathbf{X}_{-ij}$ *denotes the set of all variables except* X_{ij}.

By imposing first order symmetry on the correlation structure, we have the symmetric first order conditional autoregression (SFCAR) defined by the conditions

$$\mathbb{E}\{X_{ij}|\mathbf{X}_{-ij}\} = \frac{\lambda}{\kappa}(X_{i+1,j} + X_{i-1,j} + X_{i,j+1} + X_{i,j-1}),$$
$$Prec\{X_{ij}|\mathbf{X}_{-ij}\} = \kappa > 0,$$

where $0 \leq \lambda \leq \frac{\kappa}{4}$. Here, $\theta_{00} = \kappa$ and $\theta_{1,0} = \theta_{-1,0} = \theta_{0,1} = \theta_{0,-1} = -\lambda$. The SFCAR model is the 2-D extension of the 1-D autoregressive (AR) model that is widely used to model basic correlation in 1-D. Here in the 2-D case we have symmetric conditional dependence on four neighboring nodes in the four (planar) directions, capturing basic 2-D correlation structure. It can be shown that the GMRF defined by the SFCAR model is a zero-mean stationary Gaussian process on $\mathcal{I}_\infty$ with power spectral density [5]

$$f(\omega_1, \omega_2) = \frac{1}{4\pi^2\kappa(1 - 2\zeta\cos\omega_1 - 2\zeta\cos\omega_2)}, \quad (5)$$

where the *edge dependence factor* ζ is defined as

$$\zeta \triangleq \frac{\lambda}{\kappa}, \quad 0 \leq \zeta \leq 1/4. \quad (6)$$

The SFCAR model is useful especially because the correlation strength is captured in this single quantity ζ for SFCAR signals, which enables us to investigate the per-node information as a function of the field correlation. Here, $\zeta = 0$ corresponds to the i.i.d. case, whereas $\zeta = 1/4$ corresponds to the perfectly correlated case. Henceforth, we assume that the 2-D stochastic signal $\{X_{ij}\}$ in (1) is given by a stationary GMRF defined by the SFCAR model, as $n \to \infty$. The signal power $P \triangleq \mathbb{E}\{X_{00}\}^2 (= \mathbb{E}\{X_{ij}^2\} \, \forall \, i, j)$ is obtained using the inverse Fourier transform, and is given by $P = \frac{2K(4\zeta)}{\pi\kappa}$, $\left(0 \leq \zeta \leq \frac{1}{4}\right)$, where $K(\cdot)$ is the complete elliptic integral of the first kind [6]. Thus, the measurement SNR is given by SNR $= \frac{P}{\sigma^2} = \frac{2K(4\zeta)}{\pi\kappa\sigma^2}$.

2.1. Large System Analysis: Per-Node Information

The key idea behind the large system analysis here is that, under the stationarity assumption, the amounts of information from the node become identical regardless of sensor location as the network size grows, and the total amount of information is given approximately by the product of the number of sensor nodes and the (asymptotic) per-node information. The asymptotic per-node KLI and per-node MI are defined as

$$\mathcal{K}_s = \lim_{n\to\infty}\frac{1}{|\mathcal{I}_n|}\log\frac{p_0}{p_1}(\{Y_{ij}, ij \in \mathcal{I}_n\}) \text{ a.s. under } p_0, \text{ and}$$

$$\mathcal{I}_s = \lim_{n\to\infty}\frac{1}{|\mathcal{I}_n|}I(\{X_{ij}, ij \in \mathcal{I}_n\}; \{Y_{ij}, ij \in \mathcal{I}_n\}),$$

respectively. For the MI, the signal model (1) is applicable directly, whereas for the KLI the probability density functions of the null (noise-only) and alternative (signal-plus-noise) distributions are those given under the respective models

$$p_0(Y_{ij}) \quad : \quad Y_{ij} = W_{ij}, \quad ij \in \mathcal{I}_n,$$
$$p_1(Y_{ij}) \quad : \quad Y_{ij} = X_{ij} + W_{ij}, \quad ij \in \mathcal{I}_n. \quad (7)$$

The following closed-form expressions for the asymptotic per-node information in the spectral domain have been obtained in [2] by exploiting the spectral structure of the CAR signal and the relationship between the eigenvalues of block circulant and block Toeplitz matrices representing 2-D correlation structure.

Theorem 1 *Under the 2-D SFCAR signal model, the asymptotic per-node KLI* $\mathcal{K}_s$ *and per-node MI* $\mathcal{I}_s$ *are given by*

$$\mathcal{K}_s = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left(\frac{1}{2}\log\left(1 + \frac{SNR}{(2/\pi)K(4\zeta)(1 - 2\zeta\cos\omega_1 - 2\zeta\cos\omega_2)}\right)\right.$$
$$\left. + \frac{1}{2}\frac{1}{1 + \frac{SNR}{(2/\pi)K(4\zeta)(1-2\zeta\cos\omega_1-2\zeta\cos\omega_2)}} - \frac{1}{2}\right)d\omega_1 d\omega_2. \quad (8)$$

and

$$\mathcal{I}_s = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\frac{1}{2}\log\left(1 + \frac{SNR}{(2/\pi)K(4\zeta)(1 - 2\zeta\cos\omega_1 - 2\zeta\cos\omega_2)}\right)d\omega_1 d\omega_2,$$
$$(9)$$

respectively.

Note that the SNR and correlation are separated in (8)-(9), which enables us to investigate the effects of each term separately. With regard to $\mathcal{K}_s$ and $\mathcal{I}_s$ as functions of ζ, it is readily seen from Theorem 1 that $\mathcal{K}_s$ and $\mathcal{I}_s$ are continuously differentiable C^1 functions of the edge dependence factor ζ ($0 \leq \zeta \leq 1/4$) for a given SNR since $f : x \to K(x)$ is a continuously differentiable C^∞ function for $0 \leq x < 1$ [7]. Fig. 2 shows $\mathcal{K}_s$ as a function of ζ for several different SNR values. It is seen in the figure that at high SNR $\mathcal{K}_s$ decreases monotonically as the correlation becomes strong, i.e., $\zeta \to 1/4$. At low SNR, on the other hand, correlation is beneficial to the performance. $\mathcal{I}_s$ shows similar behaviors even if it is not shown here.

With regard to $\mathcal{K}_s$ and I_s as functions of SNR, the behavior is given by the following theorem from [2].

Theorem 2 $\mathcal{K}_s$ *and* $\mathcal{I}_s$ *are continuous and monotonically increasing as SNR increases for a given edge dependence factor* $0 \leq \zeta < 1/4$. *Moreover,* $\mathcal{K}_s$ *and* $\mathcal{I}_s$ *increase linearly with respect to* $\frac{1}{2}\log$ *SNR as SNR* $\to \infty$. *As SNR decreases to zero, on the other hand,* $\mathcal{K}_s$ *and* $\mathcal{I}_s$ *decrease to zero with convergence rates*

$$\mathcal{K}_s(SNR) = c \cdot SNR^2 + o(SNR^2), \quad (10)$$
$$\mathcal{I}_s(SNR) = c' \cdot SNR + o(SNR), \quad (11)$$

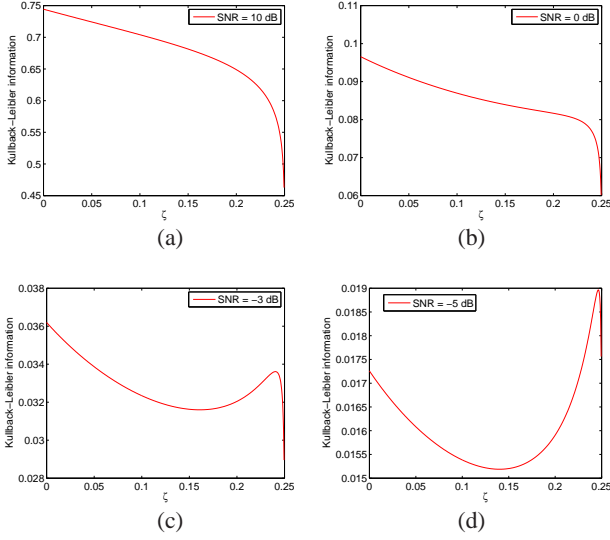*respectively, for some constants* c *and* c'.

Fig. 2. $\mathcal{K}_s$ as a function of ζ: (a) SNR = 10 dB, (b) SNR = 0 dB, (c) SNR = -3 dB, (d) SNR = -5 dB (from [2])

3. AD HOC SENSOR NETWORKS: OPTIMAL DENSITY

Based on the results in the previous sections, we now address the optimal density problem given in Section 1.

3.1. Physical correlation model

As we vary the node density for a given area with size $2L \times 2L$, the sensor spacing d_n changes. In turn, the edge dependence factor between two adjacent samples varies for given physical diffusion parameters. So, we first derive the relationship between sensor spacing d_n and the edge dependence factor ζ for the SFCAR. The physical correlation for the SFCAR model is obtained by solving the continuous-index equivalent given by the 2-D stochastic Laplace equation [8]

$$\left[\left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 - \alpha^2\right] X(x, y) = u(x, y), \quad (12)$$

where $u(x, y)$ is the 2-D white zero-mean Gaussian perturbation and $\alpha > 0$ is the diffusion rate. By solving this equation, the edge correlation factor ρ is given, as a function of the sensor spacing d_n, by [8]

$$\rho \triangleq \frac{\mathbb{E}\{X_{00}X_{10}\}}{\mathbb{E}\{X_{00}^2\}} = g(d_n) = \alpha d_n K_1(\alpha d_n), \quad (13)$$

where $K_1(\cdot)$ is the modified Bessel function of the second kind. The correlation function (13) can be regarded as the representative correlation in 2-D, similar to the exponential correlation function e^{-Ad_n} in 1-D. Both functions decrease monotonically w.r.t. d_n. However, the 2-D correlation function is flat at $d_n = 0$ [8]. Further, we have a mapping $g : \rho \rightarrow \zeta$ from the edge correlation factor ρ to the edge dependence factor ζ, given by [9]

$$\rho = \frac{(2/\pi)K(4\zeta) - 1}{4(2/\pi)\zeta K(4\zeta)} =: h^{-1}(\zeta), \quad (14)$$

which maps zero and one to zero and 1/4, respectively. Combining (13) and (14), we have a mapping $\zeta = h(g(d_n))$ from the sensor spacing d_n to ζ for the SFCAR model.

3.2. Density Analysis

The assumptions for the planar *ad hoc* sensor network that we consider is summarized in the following.

(A.1) $(2n + 1)^2$ sensors are located on the lattice $\mathcal{I}_n = [-n : 1 : n]^2$ with spacing d_n, as shown in Fig. 1, and a fusion center is located at the center $(0, 0)$. The observation samples $\{Y_{ij}\}$ at sensors form a 2-D hidden SFCAR GMRF on the lattice, and the correlation functions are given by (13) - (14).

(A.2) The fusion center collects the measurement from all nodes using minimum hop routing. A hop count of $|i| + |j|$ is required for minimum hop routing to deliver Y_{ij} to the fusion center.

(A.3) The communication energy per edge is given by $E_c(d_n) = E_0 d_n^\nu$, where $\nu \geq 2$ is the attenuation factor of wireless propagation in the physical layer.

(A.4) Sensing requires energy, and the sensing energy per node is denoted by E_s. Further, we assume that the *measurement* SNR increases linearly w.r.t. E_s, i.e., SNR $= \beta E_s$ for some constant β.[1]

The density optimization under the energy constraint can be solved using our large system analysis in the previous sections assuming the asymptotic result is still valid in low density case. The total amount I_t of information is given by

$$I_t = (2n + 1)^2 \mathcal{K}_s(\text{SNR}, d_n) \text{ or } I_t = (2n + 1)^2 \mathcal{I}_s(\text{SNR}, d_n), \quad (15)$$

for KLI or MI, respectively. The total energy E required for data collection is given by

$$\begin{aligned} E &= (2n + 1)^2 E_s + E_c(d_n) \sum_{i=-n}^{n} \sum_{j=-n}^{n} (|i| + |j|), \\ &= (2n + 1)^2 E_s + 2n(n + 1)(2n + 1)E_c(d_n). \quad (16) \end{aligned}$$

Thus, Problem 1 can be reformulated as

$$\mu_n^* = \underset{\mu_n}{\arg\max} \ (2L)^2 \mu_n \mathcal{K}_s(\text{SNR}(E, \mu_n), d_n(\mu_n)), \quad (17)$$

$$\text{s.t. } (2n + 1)^2 E_s(\mu_n) + 2n(n + 1)(2n + 1)E_c(d_n(\mu_n)) \leq E,$$

where the sensing energy E_s as well as n and d_n are functions of the node density μ_n. From $\mu_n (= (2n + 1)^2/(2L)^2)$, we first calculate n and then $d_n = L/n$. When d_n is determined, $E_c(d_n)$ is obtained from the propagation parameters E_0 and ν, and then $E_s(\mu_n)$ is obtained from the constraint in (17). Once $E_s(\mu_n)$ is determined, the measurement SNR is calculated using Assumption (A.5), i.e., SNR $= \beta E_s$ and finally we evaluate the per-sensor information $\mathcal{K}_s(\text{SNR}, \zeta(\rho(d_n)))$ and $\mathcal{I}_s(\text{SNR}, \zeta(\rho(d_n)))$ from Theorem 1.

Fig. 3 shows the total information obtainable from 2×2 square meter area as we vary the node density μ_n with a fixed total energy budget of E joules. Other parameters that we use are given by

$$\alpha = 100 \text{ (diffusion rate)}, \ \beta = 1, \ E_0 = 0.1 \text{ and } \nu = 2.$$

---

[1] Suppose that E_1 joules are required for one sensing to obtain one sample $Y_{ij}(m) = X_{ij}(m) + W_{ij}(m)$ at sensor ij and the measurement SNR of this sample is SNR_1. Now assume that we obtain M samples $(m = 1, \cdots, M)$ using M subsensors at the same location ij simultaneously, requiring $M \cdot E_1$ joules, and we take an average of these M samples at sensor ij, yielding an effective sample $Y_{ij} = (1/M) \sum_m Y_{ij}(m)$ of SNR of $M\text{SNR}_1$ assuming that the measurement noise is i.i.d. across the subsensors.
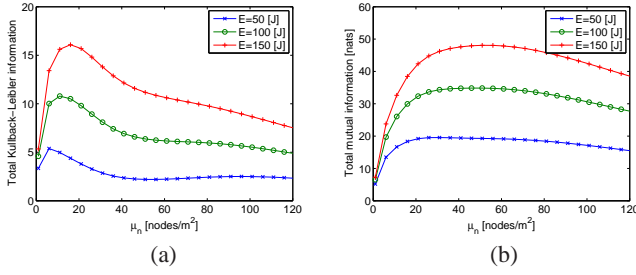
Fig. 3. (a) total KLI vs. density and (b) total MI vs. density

Here, the values of E, E_0 and β are chosen so that the minimum and maximum per-sensor sensing SNR's are roughly -10 to 10 dB for maximum and minimum densities, respectively. The diffusion rate $\alpha = 100$ is selected for the edge correlation coefficient ρ to vary from almost zero to 0.6 as the node density changes. It is seen in the figure that there is an optimal density for each value of E for both information measures. It is also seen that the total KLI is sensitive to the density change whereas the total MI is less sensitive. The existence of the optimal density is explained as follows. At low density, we have only a few sensors in the area. So, the energy for communication is not large due to the small number of communicating nodes and most of the energy is allocated to the sensing energy; the per-node sensing energy is even higher due to the small number of sensors. However, the per-node information increases only logarithmically w.r.t. the sensing energy or SNR by Theorem 2, and this logarithmic gain cannot compensate for the loss in the number of sensors. Therefore, low density yields very poor performance, and large gain is obtained initially as we increase the density from very low values as seen in Fig. 3. As we further increase the density, on the other hand, and the per-node sensing energy or SNR decreases due to the increase in the overall communication and the increase in the number of sensor nodes, and the measurement SNR is eventually at low SNR regime, where (10) and (11) hold. From (16), we have

$$E_s(\mu_n) = \beta^{-1}\text{SNR} = O(n^{-2}) \qquad (18)$$

for fixed E and $E_c = E_0(L/n)^2$, as $n \to \infty$. By the low SNR behavior of $\mathcal{K}_s$ given by (10), the behavior of the total Kullback-Leibler information is given by

$$\text{Total KLI} = (2L)^2\mu_n\mathcal{K}_s = O(n^2n^{-4}) = O(n^{-2}) = O(\mu_n^{-1})$$

and by (11) the total mutual information is given by

$$\text{Total MI} = (2L)^2\mu_n\mathcal{I}_s = O(n^2n^{-2}) = O(1).$$

This explains the initial decay after the peak in Fig. 3 (a) and quite flat curve in Fig. 3 (b). In the above equations, however, the effect of ζ on $\mathcal{K}_s$ and $\mathcal{I}_s$ is not considered. As the node density increases, the sensor spacing decreases and the edge dependence factor ζ increases for a given diffusion rate α. The behavior of the per-node information as a function of ζ is shown in Fig. 2. Note in Fig. 2 that the per-node information has a second lobe at strong correlation at low SNR while at high SNR it decreases monotonically as the correlation becomes strong. The benefit of sample correlation is evident in the low energy case ($E = 50[\text{J}]$) in 3 (a); the second peak around $\mu_n = 95$ [nodes/m^2] is observed. Note that the second peak is not so significant. Since the per-node information

decays to zero as $\zeta \to 1/4$ eventually, the total amount of information decreases eventually, as seen in the right corner of the figure, as we increases the node density.

4. CONCLUSIONS

We have considered the design of 2-D arrays of networked sensors for making inferences about 2-D correlated random fields. Under the SFCAR GMRF model, the density maximizing the total information obtainable from the network under an energy constraint has been investigated. We have seen that such an optimal density exists. At low density, the amount of information gathered is small because the logarithmic increase in the per-node information w.r.t. energy cannot compensate for the loss in the number of sensor nodes. At high density, on the other hand, the performance degrades mainly due to too much correlation between samples and low sensing energy. The optimal node density effects a trade-off between these two effects.

5. REFERENCES

[1] F. Liese and I. Vajda, "On divergence and informations in statistics and information theory," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4394-4412, Oct. 2006.

[2] Y. Sung, H. V. Poor and H. Yu, "Large deviations analysis for the detection of 2D hidden Gauss-Markov random fields using sensor networks," in *Proc. 2008 ICASSP*, Las Vegas, NY, Mar. 2008.

[3] J.-F. Chamberland and V. V. Veeravalli, "How dense should a sensor network be for detection with correlated observations?," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5099-5106, Nov. 2006.

[4] A. Anandkumar, L. Tong and A. Swami, "Detection of Gauss-Markov random field on nearest-neighbor graph," in *Proc. 2007 ICASSP*, Hawaii, USA, Apr. 2007.

[5] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applicatons*, New York: Chapman & Hall/CRC, 2005.

[6] J. Besag, "On a system of two-dimensional recurrence equations," *Journal of the Royal Statistical Society. Series B*, vol. 43, no. 3, pp. 302-309, 1981.

[7] A. Erdélyi, *Higher Transcendental Functions, Vol. II.*, New York: McGraw-Hill, 1953.

[8] P. Whittle, "On stationary processes in the plane," *Biometrika*, vol. 41, no. 3, pp. 434-449, Dec. 1954.

[9] Y. Sung, H. V. Poor and H. Yu, "How much information can one get from a wireless *ad hoc* sensor network over a correlated random field?," submitted to *IEEE Trans. Inform. Theory*, Apr. 2008.